09/693333

I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail in an envelope addressed to: Commissioner for Patents, P.O. Box 1450 Alexandria, VA 22313 on October 27, 2005.

Frank C. Eisenschenk

Frank C. Eisenschenk, Ph.D., Patent Attorney

REQUEST FOR CERTIFICATE OF CORRECTION UNDER 37 CFR 1.322 AND UNDER 37 CFR 1.323
Docket No. G-073US03REG
Patent No. 6,934,636

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicants	:	Boguslaw A. Skierczynski, Nicholas J. Schork
Issued	:	August 23, 2005
Patent No.	:	6,934,636 B1
For	:	Methods of Genetic Cluster Analysis and Uses Thereof

Mail Stop Certificate of Corrections Branch
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

REQUEST FOR CERTIFICATE OF CORRECTION UNDER 37 CFR 1.322 (OFFICE MISTAKE) AND UNDER 37 CFR 1.323 (APPLICANTS' MISTAKE)

Sir:

A Certificate of Correction (in duplicate) for the above-identified patent has been prepared and is attached hereto.

In the left-hand column below is the column and line number where errors occurred in the patent. In the right-hand column is the page and line number in the application where the correct information appears.

Patent Reads:	**Application Reads:**
Column 2, line 23:	Page 2, line 31:
"(Jin & Chalraborty,"	--(Jin & Chakraborty,--

11/01/2005 SDENBOB1 00000040 190065 6934636
01 FC:1811 100.00 DA

Column 3, line 33:	Page 4, line 19:
"said sun"	--said sum--
Patent Reads:	**Application Should Read:**
Column 6, line 9:	Page 8, line 11:
"the homozygousm"	--the homozygous--
Column 9, line 40:	Page 13, line 11:
"refers the identity"	--refers to the identity--
Patent Reads:	**Application Reads:**
Column 15, line 58:	Page 22, line 19:
"$\overline{W}_{kl} \geq \overline{Z}$"	--$W_{kl} \geq \overline{Z}$--
Patent Reads:	**Application Should Read:**
Column 16, line 41:	Page 23, line 24:
"or heterozygous in said"	--or heterozygotes in said--
Column 17, line 5:	Page 24, line 16:
"encompass anything"	--encompasses anything--
Column 21, line 4:	Page 30, line 7:
"to calculated millions, "	--to calculate millions,--.

A true and correct copy of pages 2, 4, and 22 of the specification as filed which support Applicants' assertion of the errors on the part of the Patent Office accompanies this Certificate of Correction.

The Commissioner is authorized to charge the fee of $100.00 for the amendment to Deposit Account No. 19-0065. The Commissioner is also authorized to charge any additional fees as required under 37 CFR 1.20(a) to Deposit Account No. 19-0065. Two copies of this letter are enclosed for Deposit Account authorization.

Approval of the Certificate of Correction is respectfully requested.

Respectfully submitted,

Frank C. Eisenschenk, Ph.D
Patent Attorney
Registration No. 45,332
Phone No.: 352-375-8100
Fax No.: 352-372-5800
Address: P.O. Box 142950
Gainesville, FL 32614-2950

FCE/bra

Attachments: Copy of pages 2, 4, and 22 of the specification

homogeneity assessment can be of value to any study, but depends on the identification of individuals with certain features based on some distinguishing genetic characteristics, such as forensic applications.

Analyses assessing the similarity in the genetic profiles of individuals have been pursued. For example, polymorphic microsatellites (primarily CA repeats) have been used to construct trees of human individuals that reflect their geographic origin (Bowcock *et al.*, *Nature* 368:455-457, 1994), and to study the genetic variability within and between cattle breeds (Ciampolini, *et al.*, *J. Anim. Sci.* 73:3259-3268, 1995). RFLP genotypes have been used to construct trees of individuals of different ethnicities (Mountain and Cavalli-Sforza, *Am. J. Hum. Genet.* 61:705-718, 1997). Random amplified polymorphic DNA (RAPD) markers have been used to compute genetic similarity coefficients (Lamboy, *PCR Methods and Applications* 4:31-37, 1994), and to compare phenotype and genotype in plants (Jasienski, *et al.*, *Heredity* 78:176-181, 1997).

However, these analyses often rely on a priori knowledge of the groups to which the individuals belong. Many do not permit the determination in the absence of a priori knowledge of which, and to what degree, different populations may have contributed to the genetic variation within a pool or sample of individuals. However, in the large majority of cases, individuals sampled from a population represent an "admixture" of genes from several populations. These populations are reflected in the genetic profiles of individuals and hence can defy population segregation based on traditional markers such as skin color and/or self-reported ethnic affiliation. Therefore, methods of analysis are needed to accurately determine the existence of clusters of genetically similar individuals, absent phenotypic (ethnic, for example) information. As noted previously, knowledge of the homogeneity or heterogeneity of a population can be important under many circumstances including forensics and population-based studies.

In forensics, DNA fingerprinting requires the computation of 'match probabilities' between the suspect and the DNA obtained on a victim. Match probabilities are often computed relative to a database of non-suspect DNA. The utility of the DNA contributed by non-suspects will be influenced by the amount of genetic heterogeneity among the non-suspects (Jin & Chakraborty, *Heredity* 74:274-285, 1995; Sawyer *et al.*, *Am. J. Hum. Genet.* 59:272-274, 1996; Tomsey *et al.*, *J. Forensic Sci.* 44:385-388, 1999). Thus, determining the heterogeneity of the non-suspect population

similarity for a new cluster; h) applying a non-hierarchical clustering algorithm to said ordered set of similarity data using said optimal number of clusters; i) determining the relatedness between pairs of homozygous pairs by performing a paired-pair analysis on the clusters resulting from said non-hierarchical clustering algorithm, wherein the homozygous loci of two pairs are compared pairwise to determine whether the pairs share the same homozygous alleles on the same loci; j) summing said paired-pair comparison for one pair versus all pairs in a cluster; k) computing the average sum of said paired-pair comparison for all pairs in said cluster; l) assigning values to the homozygous relatedness of each member of a pair to all homozygotes in said cluster based on whether said sum for one pair is greater than or equal to said average sum of all pairs, or whether said sum for one pair is less than said average sum for all pairs; m) comparing the number of times said sum for one pair is greater than or equal to said average sum of all pairs with the number of times said sum for one pair is less than said average sum for all pairs for each individual in said cluster; and n) dividing said cluster into a first cluster and a second cluster if there is: a first group of members of said cluster wherein said number of times said sum of one pair is less than said average sum of pairs is greater than or equal to said number of times said sum for one pair is greater than or equal to said average sum for all pairs, and a second group of members of said cluster wherein said number of times said sum of one pair is less than said average sum of pairs is less than said number of times said sum for one pair is greater than or equal to said average sum for all pairs, and wherein said first group of members are placed into said first cluster and said second group of members are placed into said second cluster.

In preferred embodiments of the invention, said traits are genetic loci and said values are assigned to said traits based on the alleles of said genetic loci. Preferably, said values are: 0 when a pair of members share no common allele; 1 when a pair of members share a common allele; and 2 when a pair of members share two common alleles. Preferably, said weights are assigned based on: sharing rare alleles between a pair of members; and sharing a homozygous genotype between a pair of members.

In other preferred embodiments, said ordered set of similarity data is present in a similarity matrix. Preferably, said similarity matrix is formed based on the pairwise similarity measure:

pairwise to determine whether the pairs share the same homozygous or heterozygous alleles on the same loci :

$$Z_{k,l} = \sum_{i=1}^{L} a_i$$

where $a_i = 1$ when two sets of pairs have the same homozygous or heterozygous alleles on the same loci, otherwise $a_i = 0$;

where L denotes the total number of loci;

where k and l each represent different pairs of individuals in a particular cluster; and where k = 1,..., N, and l = 1,..., N, where N is the number of individuals in particular cluster.

Subsequently, the average score (sum of $Z_{k,l}$) of said paired-pair comparison for all pairs in a cluster is computed:

$$W_{kl} = \sum_{i>j=1}^{N} Z_{ij}$$

$$\overline{Z} = \frac{\sum_{i>j=1(ij \neq kl)}^{N} W_{ij}}{M}$$

where M is the number of permutations of paired-pairs, and W_{kl} is the sum of the paired-pair comparison of one pair versus all pairs in a cluster.

Subsequently, the sum of the comparison of one pair versus all pairs in a cluster is compared with the average sum for all pairs in order to assign a value to the homozygous or heterozygous relatedness of each member of a pair to all homozygotes or heterozygous in a cluster: if,

$$W_{kl} \geq \overline{Z} \quad then \begin{cases} a_{i,lo} = 0 \\ a_{i,ob} = 1 \end{cases}$$

$$W_{kl} < \overline{Z} \quad then \begin{cases} a_{i,lo} = 1 \\ a_{i,ob} = 0 \end{cases}$$

for each k>l=1,...,N; and i=k,l

where a_{lo} indicates that the individual's score for W is "below" the average score for the cluster, and

where a_{ob} indicates that the individual's score for W is "above" the average score for the cluster.

UNITED STATES PATENT AND TRADEMARK OFFICE

CERTIFICATE OF CORRECTION

PATENT NO. : 6,934,636 B1

APPLICATION NO.: 09/693,333

DATED : August 23, 2005

INVENTORS : Boguslaw A. Skierczynski, Nicholas J. Schork

It is certified that errors appear in the above-identified patent and that said Letters Patent is hereby corrected as shown below:

Column 2
Line 23, "(Jin & Chalraborty," should read--(Jin & Chakraborty,--.

Column 3
Line 33, "said sun" should read --said sum--.

Column 6
Line 9, "the homozygousm" should read --the homozygous--.

Column 9
Line 40, "refers the identity" should read --refers to the identity--.

Column 15
Line 58, "$\overline{W}_{kl} \geq \overline{Z}$" should read --$W_{kl} \geq \overline{Z}$--.

Column 16
Line 41, "or heterozygous in said" should read --or heterozygotes in said--.

Column 17
Line 5, "encompass anything" should read --encompasses anything--.

Column 21
Line 4, "to calculated millions, " should read --to calculate millions,--.

MAILING ADDRESS OF SENDER:
Saliwanchik, Lloyd & Saliwanchik
P.O. Box 142950
Gainesville, FL 32614-2950

UNITED STATES PATENT AND TRADEMARK OFFICE

CERTIFICATE OF CORRECTION

PATENT NO. : 6,934,636 B1

APPLICATION NO.: 09/693,333

DATED : August 23, 2005

INVENTORS : Boguslaw A. Skierczynski, Nicholas J. Schork

It is certified that errors appear in the above-identified patent and that said Letters Patent is hereby corrected as shown below:

Column 2
Line 23, "(Jin & Chalraborty," should read--(Jin & Chakraborty,--.

Column 3
Line 33, "said sun" should read --said sum--.

Column 6
Line 9, "the homozygousm" should read --the homozygous--.

Column 9
Line 40, "refers the identity" should read --refers to the identity--.

Column 15
Line 58, "$\overline{W}_{kl} \geq \overline{Z}$" should read --$W_{kl} \geq \overline{Z}$--.

Column 16
Line 41, "or heterozygous in said" should read --or heterozygotes in said--.

Column 17
Line 5, "encompass anything" should read --encompasses anything--.

Column 21
Line 4, "to calculated millions, " should read --to calculate millions,--.

MAILING ADDRESS OF SENDER:
Saliwanchik, Lloyd & Saliwanchik
P.O. Box 142950
Gainesville, FL 32614-2950